The Marginalization Paradox

Timothy C. Wallstrom, T-13

Statistical methods of inference are increasingly important to many areas of Laboratory interest. Whether the problem is to assess the safety and reliability of the nuclear stockpile, to determine the causes of global warming, or to estimate the phylogenetic tree of the HIV virus, uncertainty comes into play and needs to be treated intelligently.

Bayesian inference has recently experienced a resurgence of interest and is heavily used in many Laboratory applications. Nevertheless, there are some striking paradoxes afflicting Bayesian inference as it is frequently practiced. In recent work, I have clarified a key paradox of Bayesian inference, the Marginalization Paradox (MP) of Dawid, Stone, and Zidek.

In the Bayesian framework, we have a *statistical model,* $p(x|\theta)$, which describes the probability of observing data x when the true parameter is θ, and a prior probability $\pi(\theta)$, which describes our uncertainty about θ before we see the data. Bayes' law allows us to infer the posterior, $\pi(\theta|x)$, which describes the new probability of θ after seeing the data:

$$\pi(\theta|x) \propto p(x|\theta)\pi(\theta).$$

The MP and some related paradoxes arise from the use of improper prior distributions, which is to say, prior distributions that integrate to infinity. Such priors arise naturally when we wish to discuss ignorance about parameters on infinite domains, such as the real line. They were used by Laplace, and are still in common use today.

The MP arises in problems in which both the data and parameter can be split into parts: $\theta = (\eta,\zeta)$ and $x = (y,z)$. For problems with appropriate symmetry, the marginal densities, $\pi(\zeta|y,z)$ and $p(z|\eta,\zeta)$, are independent of y and η, respectively; we may write them as $\tilde{\pi}(\zeta|z)$ and $\tilde{p}(z|\zeta)$, respectively. There are then two different ways to compute the marginal posterior $\pi(\zeta/z)$. It turns out that when the prior is improper, the results are often incompatible. The paradox is often dramatized as a conflict between two Bayesians, B_1 and B_2; see Fig. 1.

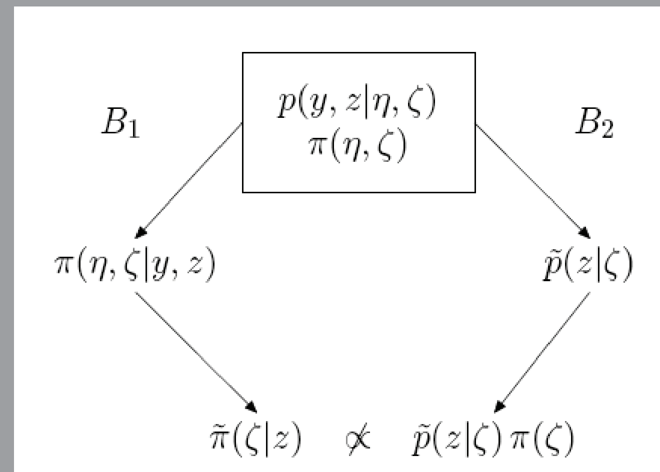

Fig. 1. B_1 and B_2 reach incompatible results for the same problem by marginalizing over different quantities.

I have studied this problem by representing the inference $\pi(\theta/\xi)$ as the limit of posteriors based on proper priors. I have shown that if the limit is taken in an appropriate sense, introduced by Mervyn Stone in the 1960s, then there is no MP.

The resolution of the MP leads to some surprising consequences. First, it turns out that many problems have no limit. This suggests that the idealization represented by the improper prior does not define a statistically meaningful problem. Second, even when the limit does exist, Bayes' law may give the wrong answer. These important consequences are under further investigation.

For more information contact Timothy Wallstrom at tcw@lanl.gov.

[1] T. C. Wallstrom, “The marginalization paradox and probability limit,” in J.M. Bernardo et al., ed., *Bayesian Statistics* **8,** 669-674, Oxford University Press (2007).
[2] T. C. Wallstrom, “The marginalization paradox and the formal Bayes’ law,” in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering,* edited by K. Knuth et al., AIP Conference Proceedings **954,** 93-100 (2007). http://arxiv.org/abs/0708.1350, and video of associated talk available on Google video (2007).

Funding Acknowledgment
- Department of Energy, National Nuclear Security Administration, Advanced Simulation and Computing